I.S.S.N: 1885-6888

ECONOMIC ANALYSIS
WORKING PAPER SERIES

Heterogeneous Focal Points, Fairness, and Coordination

♦

Raúl López Pérez

Working Paper 14/2007

UNIVERSIDAD AUTONOMA DE MADRID

DEPARTAMENTO DE ANÁLISIS ECONÓMICO: TEORÍA ECONÓMICA E HISTORIA ECONÓMICA

Heterogeneous Focal Points, Fairness, and Coordination[*]

Raúl López-Pérez[†]

August 2007

Abstract

In games with multiple equilibria, the fairest equilibrium –in case one exists- may be the obvious solution for some players but not for others, and players can be aware of this heterogeneity. This paper theoretically explores how coordination could be achieved in this case. The model is consistent with abundant experimental evidence and explains, for instance, why (a) the attractiveness of the fair equilibrium, (b) out-of-equilibrium payoffs, (c) dominated strategies, and (d) the number of players and available strategies matter for coordination. The model is compared with alternative equilibrium selection criteria like risk and payoff dominance and ideas for new experiments are suggested.

Keywords: Coordination, Equilibrium Refinement, Fairness, Focal Points, Heterogeneity.
JEL classification numbers: C72, D02, D62, D64.

---

[*] I am indebted to participants in the 2007 Pacific-Asia ESA meeting at Shanghai for helpful comments.

[†] Department of Economic Theory, Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain. E-mail: raul.lopez@uam.es

1. Introduction

Coordination becomes crucial when a group of agents have a common goal and there are different ways to achieve it. Anyone who has written a paper with a coauthor knows this well. Other examples include deciding where and when we meet somebody, selecting an industry standard, or organizing teamwork and division of labor –for a dramatic example in this line, think of a squad who must occupy a position of the enemy: A single deviation from the plan of attack can involve the utter failure of the operation, or significantly increase the number of casualties.

How do people coordinate? Game theorists often view a coordination problem as a game with multiple Nash equilibria: The players must coordinate or 'select' a particular equilibrium.[1] In this line, Schelling (1960) introduced the concept of *focal point* meaning by that any equilibrium that is 'prominent' or 'obvious' for reasons like precedence and symmetry; and argued that such prominence helps players to synchronize their expectations on their co-players' behavior. In this view, players coordinate if it is common knowledge that the game has a *unique* focal equilibrium.

In addition, Harsanyi and Selten (1988) propose two *focal principles*, or qualities that make an equilibrium focal: *Payoff dominance* -see also Gauthier (1975)- and *risk dominance* –consult Harsanyi and Selten (1988, p. 355-357) for a motivation. Equilibrium s is payoff dominant if players receive strictly higher payoffs at s than at any other equilibrium, while it is risk-dominant (in 2x2 games with two pure strategy equilibria) if it maximizes the product of players' losses from unilateral deviation.

However appealing the logic behind these two principles can be, some experimental evidence is at odds with them: Subjects sometimes fail to coordinate on the payoff dominant or the risk dominant equilibria –Cooper et al. (1990), Van Huyck et al. (1990), Straub (1995), Haruvy and Stahl (2007).

With this in mind, this paper proposes an alternative focal principle, based on one simple idea: Any equilibrium that attains a just or fair outcome is prominent or focal.[2] To explore the empirical relevance of this intuition, we follow an incremental design. Thus, section 2 offers a toy model that applies on normal-form games and is based on the hypothesis that all

---

[1] In fact, the term *coordination game* sometimes refers to a game with several equilibria, a usage that we follow throughout this paper. There is apparently no convention on this, though. Thus, Schelling (1960) uses this expression to mean any game with multiple equilibria, *all them yielding identical payoffs*; and Cooper et al. (1990) to refer to games which exhibit multiple Nash equilibria *which are Pareto-rankable*.

[2] In this paper, we use the terms 'distributive justice' and 'fairness' indistinguishably. The word fairness, in particular, does not refer exclusively to equity.

agents agree on their distributive judgments, represented by a social welfare function (SWF). This implies that players should coordinate on the unique 'fair equilibrium', *in case one exists*. To test this prediction, we consider alternative SWFs and conclude that a *Paretian and symmetric* SWF combining efficiency and equity (the *E-SWF*) fits the *available* experimental data comparatively well –Myerson (1997, p. 112) note too that "welfare properties of *equity* and *efficiency* may determine the focal equilibrium in a game" (italics in the original).

This toy model makes two remarkable predictions. First, the symmetry of the E-SWF implies that a game will not have a unique focal point (implicitly indicating that coordination is uncertain) when equilibrium payoff vectors are permutations of each other. Game 1 provides an illustration, as both (row, column) equilibrium allocations (5, 3) and (3, 5) are fair according to the E-SWF. Second, in a game with multiple surplus-preserving equilibria, players should coordinate on the most equitable one –e.g., in equilibrium (E, E) of Game 2. Yet the model is incoherent with a widely replicated experimental phenomenon: Variations in out of equilibrium payoffs often affect coordination. Thus, Schmidt et al. (2003) report a significantly higher rate of choice of action A in Game 3 than in Game 4 –see also Cooper et al. (1990), and Straub (1995).

	E1	**E2**
E1	5, 3	0, 0
E2	0, 0	3, 5

Game 1: Battle of the Sexes

	E	**A**
E	5, 5	0, 0
A	0, 0	9, 1

Game 2

	E	**A**
E	10, 10	2, 8
A	8, 2	8, 8

Game 3

	E	**A**
E	10, 10	6, 8
A	8, 6	8, 8

Game 4

The fact that subjects often choose A in Game 3 seems to indicate that they doubt that others will play E. This leads us to reconsider the hypothesis of a *common* focal point, which is arguably unrealistic in our context: People do not always coincide on what they deem fair, or simply some people do not find obvious at all a fair equilibrium.[3] In view of this, section 3 introduces *heterogeneous* focal points. More precisely, we posit that some players find obvious

[3] There seems to be few controlled evidence on this issue. However, abundant evidence indicates that people have heterogeneous social attitudes and preferences -consult Camerer (2003) or Fehr and Schmidt (2006) for surveys.

any 'E-equilibrium' leading to an allocation maximizing the E-SWF, while others do not regard fairness a focal principle.

Consider then a normal-form game with a unique E-equilibrium *s*, like Games 2, 3, and 4. If people differ beforehand on what they consider a prominent solution, how do they coordinate on *s*? We conjecture that the existence of a sufficiently large mass of E-types will be a sufficient condition for success. More precisely, we posit *s* to be focal *for all players* if the proportion of E-types is so large that no deviation from *s* is profitable when the E-players follow *s, even if the non-E-types were to uniformly deviate from s*. To put it like that, the E-equilibrium is obvious for everybody if there exists a critical mass of E-types that makes play of *s* not risky. We call this hypothesis the *E-principle*.

One can think of at least two stories behind this selection principle. First, it might be common knowledge that the E-types go always for the E-equilibrium if they believe to be numerous enough, with the result that the remaining group of people who do *not* find *s* obvious are 'forced' to play *s*, and hence everybody coordinates on *s* -this has strong similarities with the tracing procedure of Harsanyi and Selten (1988). Second, and though our model posits equilibrium behavior, it is also tempting to see the E-principle as related to recent theories of bounded rationality that assume that people have different depths of reasoning –see Camerer (2003, pp. 242-259) for a good discussion on these models.

For example, Stahl and Wilson (1995) assume that due to insufficient reason some players choose with uniform probability among their strategies (level-0 rationality), while level-k players best respond to level-(k-1) choices. From their experimental data they estimate roughly equal numbers of level-0 and level-1 agents, an insignificant fraction of level-2 subjects, and significant proportions of variously sophisticated Nash types. Along these lines, one might think of the E-types as sophisticated Nash agents who *believe* that non-E-types are basically level-0 players, and who play the E-equilibrium for sure if they are numerous enough.[4]

We offer in section 3 a number of testable predictions. Thus, the more attractive the E-equilibrium is, the easier it is to coordinate on it –e.g., consider a game identical to Game 4 except that both players get 1000 if they play (E, E). The model also predicts that out of equilibrium payoffs (*even if they correspond to dominated strategies*) and the sheer number of players and available strategies may affect coordination on the E-equilibrium. In addition, the minimal mass of E-types required to make focal the E-equilibrium *s* of a game can be used as a

[4] While we find this interpretation instructive, it is imperative to keep in mind the differences between our model and any model of bounded rationality. To put it like that, our model aims to explain towards which equilibrium play will converge, not the evolution of play.

measure of how difficult it is to coordinate on *s* on that game, and thus elaborate a ranking of games. In general, the model provides numerous ideas for new experiments –e.g., it suggests studying more thoroughly a*symmetric* games where a conflict between equity and efficiency exists.[5]

Section 3 also compares the E-principle to some alternative focal principles, and we judge it to be relatively more consistent with the evidence here analyzed. Thus, payoff dominance is at odds with the fact that out of equilibrium payoffs influence coordination, and some experimental evidence –Straub (1995), Haruvy and Stahl (2007)- attests that subjects need not always coordinate on the risk dominant equilibrium, *in particular* if it does not coincide with the E-equilibrium.

To sum up, the paper is organized as follows. Section 2 presents a toy model assuming a common focal point and considers its limitations. In section 3 we introduce heterogeneous focal points, explore its behavioral predictions, and show this expanded model to be consistent with much experimental evidence. Finally, section 4 concludes by mentioning some possible extensions.

2. A Toy Model: Homogeneous Players

Consider any two-player, normal-form game. Let $a = (ar, ac)$ denote a pure strategy profile where the row player *r* moves *ar* and the column player *c* chooses *ac*, $x(a) = [xr(a), xc(a)]$ denote its associated vector of *monetary* payoffs, and $N=\{r, c\}$ denote the set of players. To keep the model as simple as possible, we posit that *both* players are egoist and risk-neutral so that their utility function is $u_i(a) = x_i(a)$ $(i \in N)$.

In addition, we assume that the game has a *finite* number of *pure strategy* Nash equilibria and introduce a *continuous* social welfare function (SWF) $W^* : \Re^2 \to \Re$. Under these conditions, there is at least one equilibrium a^* such that $x(a^*)$ maximizes W* among all *pure equilibria* of the game.[6] We call any such a^* a W*-*focal equilibrium* and assume

Hypothesis 1: If the game has a unique W*-focal equilibrium, players follow it.

One possible interpretation of this hypothesis is that players agree on their distributional value judgments (represented by W*), and that they deem obvious any equilibrium leading to what, according to their view, constitutes the fairest possible outcome. We say that players *coordinate* if the game has a unique W*-focal equilibrium. Otherwise we simply contend that the model is undetermined (implicitly meaning that coordination failure might occur).

2.1 Testing the Model for Different Social Welfare Functions W*

[5] Much experimental analysis to date has focused on symmetric coordination games.

[6] We focus for parsimony on pure strategies though formal definitions allow for mixed ones.

How empirically relevant is this toy model? Answering this question requires being precise about the SWF W* -if we were very vague on this, any behavior could be consistent with the model and hence it would not be falsable. The problem of course is that one can think of infinite continuous SWFs. However, Konow (2003) review abundant empirical evidence on impartial fairness preferences, and conclude that people's views on fairness often integrate three principles (*efficiency*, or maximizing the surplus; *equity*, based on proportionality; and *need*, or equal satisfaction of basic needs). For this reason, but also for parsimony, we conjecture that a reasonable W* should linearly combine the *utilitarian*, *maximin and egalitarian*[7] SWFs, respectively:

$$W^u(x) = \sum_{i \in N} x_i \tag{1}$$

$$W^m(x) = \min_{i \in N} \{x_i\} \tag{2}$$

$$W^e(x) = \min_{i \in N} \{x_i\} - \max_{i \in N} \{x_i\} \tag{3}$$

To informally explore which linear combination has the largest explanatory power, we start by considering the most parsimonious candidates, that is, the very SWFs (1) to (3). In this regard, we note first that the available evidence is at odds with the assumption that W* is the utilitarian or the maximin SWFs. To illustrate this point, consider Game 5 (the so-called *Stag Hunt* game), where $a > b > 0$. This game has two equilibria in pure strategies, that is (E, E) and (S, S), but only (E, E) leads to a Pareto efficient outcome. For this reason, the toy model predicts that all players should choose E if W* were the utilitarian or the maximin SWF (or any linear combination of them).[8]

However, and as the experimental evidence from Cooper et al. (1992), Straub (1995), and Clark et al. (2001) attests, subjects often fail to coordinate on the efficient equilibrium. In the experiment by Cooper et al. (1992), for instance, subjects played a Stag Hunt game twenty times with random, anonymous re-matching, and the reported data from the last eleven periods shows that of the 330 total plays *only* 5 were of the efficient strategy.

[7] By computing the payoff distance between the worst-off and the best-off players, the egalitarian SWF provides a crude measure of the inequality embodied in distribution x. This function is not differentiable at some points, but that does not seem to pose a big problem, at least for the games we analyze. Of course, one may think of alternative, more sophisticated SWFs on this line.

[8] In contrast, the model would be undetermined if W* were the egalitarian SWF, as any pure strategy equilibria leads to an egalitarian allocation.

	E	S
E	a , a	0, b
S	b, 0	b, b

Game 5: The Stag Hunt Game

	E	B
E	a , a	0, 0
B	0, 0	b, b

Game 6

Why do subjects deviate from the efficient equilibrium? To think more about this problem it is worthy to analyze Game 6, which is identical to Stag Hunt (hence, $a > b > 0$) *except* that out-of-equilibrium payoffs are all zero. As Stag Hunt, Game 6 has two Pareto-ranked pure strategy equilibria –i.e., (E, E) and (B, B)-, of which only (E, E) is focal according to the utilitarian SWF or the maximin SWF. Therefore, both players should play E if W* were any of those SWFs (or a combination of them).

Is this prediction right? Van Huyck et al. (1992) use a game similar to Game 6 –the only difference was that players had available three strategies and there were accordingly three (Pareto ranked) equilibria- and report that *97 percent* of the subjects played the socially efficient equilibrium. Apparently, therefore, out of equilibrium payoffs are the reason why subjects deviate from the efficient equilibrium in the Stag Hunt. Clearly, the toy model cannot account for this phenomenon, *whatever the SWF we choose*. Note incidentally that neither can the payoff dominance criterion proposed by Gauthier (1975) and Harsanyi and Selten (1988).

The utilitarian SWF (and the payoff dominance criterion too) presents another problem. To illustrate it, consider the so-called *Nash bargaining game*: Two players simultaneously demand some amount of money between \$0 and, say, \$10 and their respective demands are satisfied if they add up to \$10 or less, while both players get \$0 otherwise. Any division of the ten dollars is clearly a surplus-preserving Nash equilibrium, so that our model is undetermined if W* is the utilitarian SWF. However, the evidence from Roth and Malouf (1979) indicates that the egalitarian sharing (5, 5) has a prominence that makes it the obvious solution. In addition, 82 percent of the participants in one game in Van Huyck et al. (1992) played the equilibrium yielding equal payoffs (5, 5) instead of two other equilibria yielding payoffs (7, 3) and (3, 7), respectively.

This problem of the utilitarian SWF leads us naturally to consider the egalitarian SWF. Nonetheless, this SWF also appears to have important problems in explaining coordination. To start, the toy model predicts coordination on equilibrium (B, B) in Game 7 because (10, 50) is the most egalitarian equilibrium outcome. However, the evidence from Straub (1995) is clearly at odds with that prediction. In this experiment, 10 subjects played this game 9 times with anonymous re-matching, and the author reports that most subjects played E most (if not all) of the time.

	E	B
E	90, 40	0, 0
B	0, 0	10, 50

Game 7

	E	A
E	100, 80	0, 0
A	0, 0	1, 1

Game 8

	E	B
E	100, 100	1,0
B	0, 1	25, 25

Game 9

Apparently, the most equitable equilibrium is not necessarily the focal one. To further pursue this point, consider Game 8. The toy model predicts coordination on equilibrium (A, A) if W* is the egalitarian SWF. Nevertheless, introspection indicates that (E, E) is much more likely. Incidentally, this problem is likely to be shared by any other W* that, as the egalitarian SWF, is not increasing –i.e., by any SWF that does not satisfy the *Paretian* property, which entails two conditions: (a) If $u'_i \geq u_i \quad \forall i \Rightarrow W^*(u') \geq W^*(u)$, and (b) if $u'_i > u_i \quad \forall i \Rightarrow W^*(u') > W^*(u)$. In addition, the assumption that equality makes an equilibrium focal implies that the toy model is undetermined in Game 9 -both (E, E) and (B, B) lead to an equal outcome. That seems a rather poor forecast, though, at least when compared with the prediction that players coordinate on (E, E) if W* is the utilitarian or the maximin SWF.

The previous discussion apparently suggests that an equitable equilibrium is focal when it is socially efficient as well (as in the Nash bargaining game), but not otherwise. This insight leads us to consider a linear combination of the SWFs (1) and (3), the *efficiency-equity (E) SWF* ($0 < \delta < 1$):

$$W^E(x) = \sum_{i \in N} x_i + \delta \left(\min_{i \in N} \{x_i\} - \max_{i \in N} \{x_i\} \right) \tag{4}$$

The interpretation of hypothesis $0 < \delta < 1$ is that, when evaluating the fairness of different payoff vectors, players weight social efficiency more heavily than equality –as a result, function (4) is a *strictly* Paretian SWF so that any payoff-dominant equilibrium payoff vector maximizes it.

Consistent with the available evidence from the lab, the E-SWF predicts coordination on the surplus-maximizing outcome of Games 6 and 7 (and on the egalitarian equilibrium of the Nash Bargaining game). Furthermore, it shares the reasonable predictions of the utilitarian SWF in Games 8 and 9 and, although it cannot explain why out of equilibrium payoffs affect coordination, we have seen that this problem afflicts any SWF in our simple model.

For all these reasons, we find the E-SWF a particularly appealing choice. Of course, one might think of other sensible alternatives, like a linear combination of the utilitarian and the maximin SWFs.[9] Nevertheless, in two-player games this latter SWF makes the same predictions as the E-SWF, and we have scant lab evidence from multi-player, one shot, coordination games in which equity and maximin conflict. Game 10 is an extremely simple illustration of this class of games: Three players (Row, Column and Dummy) are involved but only Row and Column are active, while Dummy has no say –in each cell, Dummy's payoff is the right-hand one.

	M	E
M	(60, 15, 15)	(0, 0, 0)
E	(0, 0, 0)	(40, 40, 10)

Game 10: A conflict between maximin and equity

Observe that (60, 15, 15) is W*-focal according to any SWF that linearly combines the utilitarian and maximin SWFs, while (40, 40, 10) is focal according to the E-SWF –note that both allocations are surplus-preserving. Hence, players should coordinate on the equilibrium (M, M) if they find focal the first SWF and on equilibrium (E, E) if they find focal the E-SWF. Controlled experimental evidence on this line would be welcome.

2.2 Symmetric Social Welfare Functions: Some implications

It is worthy to remark that all examples of W* that we have considered up to now satisfy the property of *symmetry*, that is, W*(x) = W*(x') if the entries of payoff vector x constitute a permutation of the entries of vector x'. This means that in evaluating the fairness of an outcome, all agents are on the same footing. Since most of us probably regard *impartiality* as a minimal requirement on any justice theory –Barry (1995)-, we find symmetry a reasonable property, at least in one-shot games played by anonymous subjects, as is usual in many experiments.

To reflect on the consequences of choosing a symmetric W*, consider any game with at least two *equilibrium* payoff vectors x, x' that constitute a permutation of each other *and* maximize W*. Clearly, the toy model does not select a unique equilibrium and hence it is undetermined here. We understand this multiplicity of predictions as a signal that real players may fail to coordinate, something that appears to be consistent with the available experimental data.

As an illustration, consider the 'Battle of the Sexes' (BOS) Game 11, where $a > b > 0$. This game has two Nash equilibria in pure strategies –i.e., (E1, E1) and (E2, E2)- and since the associated equilibrium payoff vectors are permutations of each other, the toy model is clearly undetermined if W* is symmetric**.** Consistent with this indeterminacy, Cooper *et al.* (1989) and

[9] In view of the experimental evidence shown by Konow (2003), one could also think of a more

Straub (1995) report experimental evidence from this game and show that subjects often fail to coordinate on any pure strategy equilibrium. Although rather erratic, behavior roughly matches the mixed strategy equilibrium prediction as subjects tend to choose their preferred move with the highest probability. Interestingly, however, Straub (1995) make clear that his results fail to replicate those from Cooper et al (1989) –in fact, Straub ran two sessions with different subjects but an identical BOS, and reported significantly different results from session to session.

	E1	**E2**
E1	a , b	0 , 0
E2	0 , 0	b , a

Game 11: The Battle of the Sexes

	E1	**E2**
E1	a , a	0 , 0
E2	0 , 0	a , a

Game 12: Matching game

Game 12 ($a > 0$) provides further illustration. This is an example of a *matching game* -Camerer 2003, p. 341- because all equilibria have the same payoffs for each player. As its two pure strategy equilibria are W*-focal according to any symmetric SWF, it follows that the toy model is again undetermined. According to the model, therefore, coordination fails in BOS and matching games for the same reason: The lack of a unique E-equilibrium or, in other words, a unique 'fair manner' to play the game. This is in contrast with an alternative argument sometimes given to explain coordination failures in BOS, that is, that subjects have in general doubts about which one deserves the better outcome. We do not believe this to be right, though, as it is at odds with the above mentioned evidence from Game 7: Apparently, subjects have no doubt about which player deserves the better outcome *if there is clearly a fair outcome*.

Incidentally, we would like to stress that players may have available other focal points in matching games or BOS so that coordination need not be simply a matter of luck. Alternative focal principles include precedence, the strategies' labels, the uniqueness of a certain physical property, etc. For instance, 89 percent of the participants in one experiment reported in Mehta, Starmer and Sugden (1994) coordinated on Mt. Everest when choosing from a large set of mountains in a matching game.

This raises an interesting question: Do subjects attach preference to any particular focal principle when some of them collide? For instance, consider a game in which two *Chinese* players simultaneously choose between 'Mt. Everest' or 'Mt. Teide' (the highest elevation in Spain), and such that each player earns \$5 if they coordinate on 'Mt. Everest', \$10 if they coordinate on 'Mt. Teide', and no money otherwise. Clearly, the choice 'Teide' stands from the distributive point of view, whereas 'Everest' stands for cultural reasons.

complex SWF including social efficiency, equity, and the maximin concern as ingredients.

As another example, participants in one of the treatments of Van Huyck et al. (1992) were recommended (‘assigned’, in the authors’ terminology) some equilibrium in a game similar to our Game 6. When the payoff-dominant equilibrium was recommended, subjects unanimously played it; however, when the payoff-dominated equilibrium was assigned, less than half of the subjects played it. This suggests that, as a focal point, fairness is more ‘prominent’ than assignments, at least for a majority of the subjects.

2.3 Conclusions

On one hand, the hypothesis that W* is symmetric and depends on both efficiency and equity produces two appealing predictions: (a) Coordination is uncertain if the equilibrium outcomes are permutations of each other, and (b) both the social surplus and payoff disparity matter for coordination. On the other hand, the toy model fails to explain why out of equilibrium payoffs affect coordination. In order to account for this, the following section extends the model while keeping the E-SWF as a key ingredient.

3. Extending the Model: Heterogeneous Agents

We hence relax the hypothesis of homogeneity, positing instead that players have heterogeneous views on what they consider a focal equilibrium and that, moreover, each player is uncertain about each other player’s view. For this, we assume that players play a Bayesian game. That is, prior to the start of the game, players’ types are assumed to be drawn from some objective distribution over a finite set $W = \{W_1, W_2, .., W_K\}$ of K *continuous and symmetric* social welfare functions. More precisely, and to simplify the analysis, we consider only two types (the model can be extended to more complex cases):

Hypothesis 2: The type of a proportion ρ of the players is the E-SWF ($0 < \delta < 1$)

$$W^E(x) = \sum_{i \in N} x_i + \delta\left(\min_{i \in N}\{x_i\} - \max_{i \in N}\{x_i\}\right) \tag{4}$$

Remaining players’ type is the ‘constant’ SWF: $W^C(x) = C \qquad (C \in \Re)$. ■

Intuitively, players whose type is the constant SWF do find all equilibria outcomes equally prominent. For them, therefore, fairness does not play any role in equilibrium selection. On the contrary, an equilibrium outcome is prominent for the other players (the *E-types*) if it maximizes the E-SWF.[10] Any W^E-focal equilibrium of a game will be called an *E-equilibrium* hereafter.

If $\rho = 1$ the model collapses to the toy model plus the assumption that W* is the E-SWF. In this case, players follow the E-equilibrium in case only one exists. To analyze the case

[10] Note well that types do not differ in their utility function (both are risk neutral and selfish) but in their beliefs about what a prominent equilibrium is. In fact, one could depict the ideas of the model without making reference to Bayesian games, but at the cost, we believe, of expositional clarity.

$\rho < 1$, consider any game with a *unique* E-equilibrium $a_E = (a_{Er}, a_{Ec})$, and let m_i denote the number of *pure* strategies available to player *i*, and *ai'* denote any *non-dominated* pure strategy different than a_{Ei} (*i* = *{r, c}*). Given this, we define threshold $\rho_{Er} \in [0,1]$ as the minimum ρ such that any $\rho \geq \rho_{Er}$ satisfies

$$\rho \cdot x_r(a_E) + \frac{1-\rho}{m_c - 1} \sum_{a_c \neq a_{Ec}} x_r(a_{Er}, a_c) \geq \max_{a_r'} \left[\rho \cdot x_r(a_r', a_{Ec}) + \frac{1-\rho}{m_c - 1} \sum_{a_c \neq a_{Ec}} x_r(a_r', a_c) \right]. \quad (5)$$

In other words, when $\rho \geq \rho_{Er}$ the row player does not find optimal to deviate from a_{Er} if she expects the column player to choose a_{Ec} if he is an E-type and *any pure strategy different than* a_{Ec} with equal probability if he is not an E-type. The computation of ρ_{Er}, which obviously takes a specific value in each game, is greatly simplified by noting that the optimal *ar'* in the right-hand side of (5) has necessarily to be a best response to a_{Ec} among all strategies different than a_{Er}, since it must maximize the right-hand side of (5) in particular when $\rho = 1$.[11] The column player's threshold ρ_{Ec} is analogously defined and is equal to ρ_{Er} if the game is symmetric. In turn, we define ρ_E as $\max\{\rho_{Er}, \rho_{Ec}, 0\}$. Note that ρ_E always exists since we allow for the possibility that $\rho_E = 1$. This concept leads to the following

Hypothesis 3 (the E-Focal Principle): In games with a unique E-equilibrium, players follow it if $\rho \geq \rho_E$.[12] No focal equilibrium exists otherwise.

The model can be easily applied to two-player coordination games with 2x2 matrices, like the one at Table 1. To make this a coordination game with just one E-equilibrium, assume without loss of generality that (E, E) and (NE, NE) are the only pure strategy equilibria and that (E, E) is the E-equilibrium.

[11] Since *ar'* is restricted to be pure, there is always such an optimal deviation. With mixed strategies, in contrast, an optimum need not exist –the problem is that a deviation may become more profitable as the mixture gets closer to a_{Ei}; game 16 of Haruvy and Stahl (2007) is an example of this.

[12] This hypothesis is silent about games like BOS, matching games or indeed any game with multiple E-equilibria. Implicitly, the model is undetermined here.

	E	NE
E	a_{11}, b_{11}	a_{12}, b_{12}
NE	a_{21}, b_{21}	a_{22}, b_{22}

Table 1: A 2 x 2 Payoff Matrix

One can show that ρ_E coincides in this game with the highest possible probability that a player chooses the E-action in the mixed strategy equilibrium. In effect, application of equation (5) to Table 1 gives us the following value for ρ_{Er} :

$$\rho \cdot a_{11} + (1-\rho) \cdot a_{12} \geq \rho \cdot a_{21} + (1-\rho) \cdot a_{22}$$

$$\Leftrightarrow \rho \geq \frac{a_{22} - a_{12}}{a_{22} - a_{12} + a_{11} - a_{21}} = \rho_{Er}, \qquad (6)$$

that is, equal to the probability that the row player chooses E in the mixed strategy equilibrium of the game. In addition, an analogous argument shows that the threshold prior for the column player is

$$\frac{b_{22} - b_{21}}{b_{22} - b_{21} + b_{11} - b_{12}} = \rho_{Ec} \cdot \qquad (7)$$

Since ρ_E coincides by definition with the maximum of (6) and (7), this proves our prior statement. As an illustration, direct application of equations (6) and (7) confirm that ρ_E is equal to *b/ a* in the Stag Hunt Game 5 and to *b/(a + b)* in Game 6.

Inspection of (6) and (7) provide one key insight. Recall that hypothesis 3 says that players will play E for sure if the proportion of E-types is larger than ρ_E. For this reason, coordination on the E-equilibrium should become easier as thresholds (6) and (7) decrease. In this regard, partial derivation demonstrates that ρ_{Er} increases with a_{22} and a_{21} and decreases with a_{11} and a_{12} -an analogous pattern holds for ρ_{Ec}. Therefore, *both equilibrium and out of equilibrium payoffs are crucial to ensure coordination*.

We believe that this insight is supported by much experimental evidence. To start, it might explain why subjects coordinate on the E-equilibrium of Game 6 much more frequently than in the Stag Hunt Game 5 –we have just seen that the threshold associated to Game 6 is smaller than that of Game 5. As further favoring evidence, Straub (1995) study Games 13 and 14 (points indicate the probability of earning $2), and report changes in efficient play in a direction coherent with our model. Different cohorts of subjects played each game 9 rounds and while play converged to (E, E) in Game 13, it converged to equilibrium (NE, NE) in Game 14. Observe that ρ should be at least equal to 0.5 and 0.75 for players to coordinate on (E, E) in Game 13 and 14, respectively –see also Battalio et al. (2001) and Schmidt et al. (2003) for similar results.

In addition, Straub (1995) studies a variant of Game 14 where a_{12} and b_{21} were equal to 0 instead of 20, and reports convergence to equilibrium (NE, NE) –according to our model, ρ should be at least equal to 0.8 to ensure coordination on (E, E). Interestingly, among the 6 games with a unique E-equilibrium that Straub analyzed, players coordinated on the E-equilibrium *only* in those games which had a threshold prior ρ_E *smaller or equal* than 2/3.

	E	**NE**
E	100, 100	20, 60
NE	60, 20	60, 60

Game 13

	E	**NE**
E	100, 100	20, 80
NE	80, 20	80, 80

Game 14

In another multi-game study, Haruvy and Stahl (2007) consider 14 symmetric, 3x3 payoff matrices with two or three Pareto-ranked equilibria. Remarkably, the E-equilibrium was played by the majority of the subjects *only* in those games where the critical mass ρ_E was *smaller or equal than 0.27*. In other words, our model correctly partitions the set of 14 games between one subset where players coordinate (on the E-equilibrium) and one subset where players do not. Nevertheless, the limit number 0.27 contrasts with the aforementioned limit of 2/3 from Straub (1995).[13]

As we noted, equilibrium payoffs (both non-E and E-equilibrium ones) should also influence coordination. More precisely, our model makes the following testable prediction: Ceteris paribus, coordination on the E-equilibrium of Table 1 becomes easier as its associated surplus (or attractiveness) $a_{11} + b_{11}$ increases. To illustrate this point, consider a variant of Game 14 where a_{11} , b_{11} are both equal to 1000 instead of 100: It *seems* intuitive that coordination is more likely in the former case. Analogously, the *less* attractive the non-E-equilibrium is, the easier is to coordinate on the E-equilibrium.

There is some controlled evidence on this specific issue, although we also find that additional study could be worthwhile. Brandts and Cooper (2007) study coordination in a class of games more complex than the ones we consider here and observe a higher level of

[13] Subjects in Haruvy and Stahl (2007) were encouraged to use calculators in the main treatment and, as the results from a control treatment indicated, this feature induced more level-1 choices and reduced significantly the play of the E-equilibrium in most of the games analyzed. Further, the payoff matrices that subjects viewed in their screens displayed only *their own* payoffs. The games being symmetric, this should be inconsequential for our model, but maybe it attenuated the prominence of any equilibrium in terms of fairness. We speculate that these two factors reduced notably subjects' beliefs about ρ , thus explaining the quantitative discrepancy between both studies.

coordination on the E-equilibrium as it becomes more attractive.[14] In turn, Schmidt et al. (2003) consider an array of *symmetric* 2x2 games with a unique payoff dominant equilibrium and different values of what they call payoff dominance measure -if the game at Table 1 is supposed to be symmetric, this measure is computed as the ratio P= $a_{11} - a_{22} / a_{11}$. However, *no* pair of games in their study differed *only* in the attractiveness of the E-equilibrium.

It is worthy to remark in this regard that our model predicts that *ceteris paribus* an increase in P should facilitate coordination on the E-equilibrium, because either decreasing a_{22} or increasing a_{11} should foster coordination. In contrast, Schmidt et al (2003) claim from their results that changes in the level of P do not affect significantly subjects' play. Nevertheless, it is unclear from our point of view what drove their results because they simultaneously changed out of equilibrium payoffs. In our view, more experimental evidence on this issue is required.

To finish, note well that hypothesis 3 selects the E-equilibrium as the obvious solution *only* if the mass of E-players is 'large enough'. For instance, the model *unequivocally* predicts that Stag Hunt players will move E if *a* is sufficiently *larger* than *b*, (this is one way to rephrase the above mentioned condition $\rho \geq b/a = \rho_E$) but it is undetermined otherwise. We interpret this indetermination as a signal that players may fail to coordinate and that multiple outcomes are possible. Consistent with this, different experimental studies report *different* rates of efficient play when *a* and *b* are 'close enough'. In Cooper et al. (1992), payoffs (*b* =800; *a* =1000) were given in points that determined the probability of the player winning a lottery where winning players received $1 and losing players received $0 -thus, *a* and *b* were arguably close. The reported data shows that play of the *inefficient* equilibrium was prevalent –Clark et al. (2001) replicate this result. In contrast, Duffy and Feltovich (2002) show much larger levels of efficient play *even* when they used a similar payoff calibration and binary lottery procedure. Moreover, they report a large variance in the three sessions that they ran: In two of them, the frequency of efficient play was close to 50%, while in the other one it was 81%.

3.1 How the E-principle compares with Other Focal Principles

It may be instructive to compare our equilibrium refinement with others that have been proposed in the literature. For instance, Harsanyi and Selten (1988) propose risk dominance and payoff dominance. To recall, equilibrium *s* is payoff dominant if players receive higher payoffs at *s* than at any other equilibrium, while it is risk-dominant if it maximizes the product of the losses from unilateral deviation –this definition specifically applies to 2x2 payoff matrices with two pure equilibria; for a more general definition, see Haruvy and Stahl (2007). In the Stag

[14] They also observe that communication is more efficiency-enhancing than financial incentives.

Hunt Game 5, for instance, equilibrium (E, E) is payoff dominant, while equilibrium (S, S) is risk-dominant if $2b > a$ because in that case we have $(b-0)\cdot(b-0) > (a-b)\cdot(a-b)$.

Two differences between payoff dominance and the E-principle stand out, that is, the former criterion does not take into account (i) out of equilibrium payoffs and (ii) equity. With regard to risk dominance, however, the differences are much more subtle. It is a remarkable illustration of this that the data from Straub (1995), which is largely consistent with our model, is also much consistent with risk dominance: Subjects converged in *most* games in Straub (1995) to the risk-dominant equilibrium – however, risk dominance fails to explain much results in Haruvy and Stahl (2007).

Indeed there is a close relation between risk dominance and the E-principle. To see this in some detail, consider again Table 1. Let $L_r^E = a_{11} - a_{21}$ and $L_c^E = b_{11} - b_{12}$ denote the row and column player's loss from unilaterally deviating from the E-equilibrium, respectively - L_r^{NE} , L_c^{NE} are analogously defined. Then thresholds (6) and (7) can be respectively rewritten as

$$\frac{L_r^{NE}}{L_r^{NE} + L_r^E}, \qquad \frac{L_c^{NE}}{L_c^{NE} + L_c^E}.$$

We observe here, therefore, that coordination on the E-equilibrium becomes easier as L_r^E and L_c^E increase and as L_r^{NE} and L_c^{NE} decrease, that is, as the product $L_r^E \cdot L_c^E$ of the losses from unilateral deviation from equilibrium (E, E) increases and that of (NE, NE) decreases. In other words, an increase in the risk dominance level of the E-equilibrium makes more likely that it is played.

Although the relation between risk dominance and the E-principle is somehow close, it is however important to stress the disparities. First, since risk dominance does not depend on social efficiency and equity, the E-equilibrium and the risk dominant one need not coincide. As an example of an equitable equilibrium which is not risk dominant, consider Game 15. (E, E) is the E-equilibrium if the inequity parameter δ is larger than 1/3, but (NE, NE) is the risk-dominant equilibrium. Our model indicates that players should play (E, E) if ρ is at least equal to 0.625 –as application of thresholds (6) and (7) proves.

	E	**NE**
E	30, 30	0, 0
NE	0, 0	50, 20

Game 15

	E	**NE**
E	80, 80	10, 70
NE	70, 10	30, 30

Game 16

An example of a socially efficient equilibrium which is not risk dominant is equilibrium (E, E) of Game 16. Hypothesis 3 predicts that players will unequivocally play (E, E) if ρ is at least equal to 2/3. Interestingly, Straub (1995) report that participants in Game 16 predominantly choose action E.

Second, *and even if the E-equilibrium is risk dominant as well*, it is crucial to stress that hypothesis 3 selects the E-equilibrium only for 'large' values of ρ, while it makes no precise prediction otherwise -in contrast, the risk dominance criterion makes no such distinction. If players had different beliefs about ρ -i.e., if priors were not common-, this might be important. To illustrate this point, consider Game 17. Its E-equilibrium (E, E) is risk dominant and hypothesis 3 predicts play of E if both players believe that ρ is higher than 0.47 –apply equation (6) or (7). On the contrary, a subject could choose NE if she thought that ρ is lower than 0.47. Indeed, Keser and Vogt (2000) report that almost half of the subjects chose action NE and perhaps this could be explained by some heterogeneity in subjects' beliefs.[15] Incidentally, this remark points out that eliciting beliefs about ρ -e.g., by asking subjects what proportion of players find obvious an efficient and egalitarian equilibrium- could offer insightful information on how players coordinate.

Another selection principle that appears in the literature is *security* –see for instance Van Huyck et al. (1990). Roughly speaking, any equilibrium is secure if the equilibrium actions maximize each player's minimum possible payoffs (in the matrix that results after deletion of non-equilibrium actions). In contrast with the E-principle, security takes into account neither social efficiency nor equity. As a consequence, security selects equilibrium (NE, NE) in Game 18, even when play of E (which according to our model is sure if ρ is larger than 0.01) seems introspectively much more likely. In addition, the security principle does not provide a convincing explanation of why out of equilibrium payoffs affect coordination (for instance, security cannot explain why the level of efficient play differs from Game 13 to Game 14).

	E	**NE**
E	70, 70	5, 20
NE	20, 5	50, 50

Game 17

	E	**NE**
E	100, 100	0, 1
NE	1, 0	1, 1

Game 18

3.2 Dominated Strategies and Coordination

[15] Subjects in Keser and Vogt (2000) played Game 17 just once, and one could speculate that they might converge to the E-equilibrium if they played the game repeatedly and were thus allowed to learn the objective value of ρ (assuming, of course, that ρ is larger than 0.47).

Recall from expression (5) that the critical mass ρ_E characterizing hypothesis 3 is defined, for each game, as the minimal mass of E-types such that if all E-types move according to profile *a* E then no player finds optimal to deviate from *a* E, *even* if the non-E-types choose *any other strategy* with uniform probability.

We did not exclude dominated strategies from this definition. Although we believe that this issue is largely an empirical matter (and the evidence from Cooper et al. (1990) or Crawford (1997) is in favor of our approach), there are at least two arguments to justify the inclusion of dominated strategies in this definition. First, much experimental evidence proves that numerous subjects do not only care about their own material payoff –see Camerer (2003) for a survey. Consequently, a player might choose a strategy that is dominated in monetary terms –i.e., a strategy that consistently gives a lower *monetary payoff* than another one, no matter what others do- because she expects to compensate the material loss with some psychological or emotional benefit. Second, the evidence also indicates that some subjects believe that others may play dominated strategies - with frequencies ranging from 20% to just over 80%, according to Crawford (1997, p. 11).

One might think of alternative definitions of threshold ρ_E that exclusively allow for play of non-dominated strategies. To understand some of the consequences of this modeling choice consider Game 19, where strategy NE1 dominates strategy NE2, and Game 20, which is obtained from Game 19 just by deleting the dominated strategy NE2. Under hypothesis 3, the E-equilibrium (E, E) is the focal point of Game 19 if

$$\rho \cdot 60 + \frac{(1-\rho)}{2} \cdot (10+0) \geq \rho \cdot 10 + \frac{(1-\rho)}{2} \cdot (30+200),$$

that is, if ρ is higher than 0.68. In contrast, if dominated strategies are assumed not to count for the computation of the critical mass ρ_E, the E-equilibrium is focal in Game 19 if it happens to be focal in Game 20 as well, that is, if ρ is higher than 0.29 – as one can confirm by applying either (6) or (7). In other words, if dominated strategies are assumed not to matter, coordination on the E-equilibrium of Game 19 is arguably a rather easy task, while the opposite is true otherwise.

	E	**NE1**	**NE2**
E	60, 60	10, 10	0, 0
NE1	10, 10	30, 30	200, 0
NE2	0, 0	0, 200	40, 40

Game 19

	E	**NE1**
E	60, 60	10, 10
NE1	10, 10	30, 30

Game 20

3.3 Coordination May Depend on the Number of Strategies Available

One topic that seems to have received scarce systematic attention in the literature is how the sheer number of available strategies affects coordination. Given how the critical mass (5) is defined, this variable clearly plays a role in our model. We analyze its influence in a bit of detail by showing first that an increase in the number of strategies can sometimes *facilitate* coordination on the E-equilibrium. For that, consider two simplified versions of the Nash bargaining game (see 2.1). In the first one (Game 21), players can only demand \$4, \$5 or \$6, while in the second one (Game 22) they can additionally demand either \$1 or \$9.

	\$4	**\$5**	**\$6**
\$4	4, 4	4, 5	4, 6
\$5	5, 4	5, 5	0, 0
\$6	6, 4	0, 0	0, 0

Game 21

	\$1	**\$4**	**\$5**	**\$6**	**\$9**
\$1	1, 1	1, 4	1, 5	1, 6	1, 9
\$4	4, 1	4, 4	4, 5	4, 6	0, 0
\$5	5, 1	5, 4	5, 5	0, 0	0, 0
\$6	6, 1	6, 4	0, 0	0, 0	0, 0
\$9	9, 1	0, 0	0, 0	0, 0	0, 0

Game 22

Clearly, both games have a unique E-equilibrium, that is, the equal sharing. Further, direct application of expression (5) indicates that players should coordinate on the E-equilibrium in Game 21 and 22 if ρ is larger than 3/5 and 1/3, respectively. In other words, coordination is easier in Game 22, where players have five strategies available, than in Game 21, where they have only three. The intuition here is that the best possible deviation from the E-equilibrium –i.e., the \$4 demand- becomes riskier as the number of strategies increase.

This result can be easily generalized. For that, let *m* denote the number of available strategies in the Nash Bargaining game (to simplify the exposition, we consider a symmetric version of the game with an *odd* number of strategies), K the ‘E-demand’, and K-ε the largest possible demand among all demands smaller than K (K, ε > 0). Given this, expression (5) holds if

$$\rho \cdot \mathrm{K} + \frac{(1-\rho)}{m-1} \cdot \frac{(m-1)}{2} \cdot \mathrm{K} \geq \rho \cdot (\mathrm{K} - \varepsilon) + (\frac{1-\rho}{m-1}) \cdot \frac{(m+1)}{2} \cdot (\mathrm{K} - \varepsilon) \Leftrightarrow \rho \geq \frac{2\mathrm{K} - \varepsilon(m+1)}{2\mathrm{K} + \varepsilon(m-3)} = \rho_E \cdot$$

Partial differentiation then shows that the critical mass ρ_E decreases as *m* increases in the Nash Bargaining game, as we wanted to show (note also that coordination is facilitated as well if ε increases).

We want to stress that this result need not extend to other games, as a simple example can confirm. In effect, consider Game 23, where column player’s strategies NE1 and NE2 are undistinguishable. Our model predicts that both players will choose E if ρ is larger than 2/7 and, as the reader can easily confirm, this would hold as well if we added another strategy NE3 undistinguishable from NE1 and NE2. To put it like that, the addition of redundant strategies is immaterial for coordination because it does affect the level of risk associated to the play of the E-equilibrium. Interestingly, this contrasts with the prediction given by bounded rationality models of levels of reasoning (Stahl and Wilson, 1995). For instance, a level-1 row player should move E in Game 23, given her beliefs that the column player is a level-0 type who uniformly randomize among *any* strategy; *but* she should choose NE if an undistinguishable strategy NE3 was added.

	E	**NE1**	**NE2**
E	7, 7	0,0	0, 0
NE	2, 0	2, 1	2, 1

Game 23

3.4 Multiple player games

We finish this section by making a brief reference to multiple player games. Extending the model to *n*-player, normal-form games is direct as it just suffices to redefine in an obvious manner the threshold prior ρ_E so that E-types do not find optimal to deviate from the E-equilibrium if $\rho \geq \rho_E$ and any non-E-player is expected to randomize uniformly among all non-E-actions.

One class of multiple player games that have received much attention in the experimental literature are *order-statistic games*, where *n* players simultaneously choose a number within certain range, and any player's payoff depends *negatively* on the distance between her choice and a particular statistic of others' choices, like the median or the minimum. An example of the Minimum (or Weak-Link) game could be a group of people who agreed to meet at a certain place at a certain time and who can independently choose whether they arrive punctual: Everybody could prefer the punctual equilibrium, but punctual players lose if at least one person arrives late. Seminal experimental papers are Van Huyck et al. (1990, 1991) –consult Camerer (2003) and Devetag and Ortmann (forthcoming) for a review of the literature.

Order-statistic games are similar to Stag Hunt in that there is usually a tension between the efficient equilibrium and the secure one –in fact, Stag Hunt is an extremely simple version of the Minimum game. For this reason, many of the insights of our model for Stag Hunt also hold here: Coordination on the E-equilibrium is facilitated as it becomes more attractive (consistent with the evidence reported in Brandts and Cooper, 2005), and as out of equilibrium payoffs are conveniently varied (consistent with Van Huyck et al., 1990). In our prior example, everybody arrives on time if the proportion of people who find prominent the punctual equilibrium (maybe because they feel that arriving punctual is their moral duty) is ‘large enough’.

To finish, this class of games raise and interesting question, that is, should an increase in the number of players hinder coordination? To think about this issue, consider the Stag Hunt game with *n* players: Each player chooses between actions E and S; she gets a sure payoff of *b* if she moves S, a payoff of *a (a* > *b*) if all players move E, and 0 if she moves E and *at least one* player moves S.

Obviously, all players move E in the unique E-equilibrium, and hypothesis 3 indicates that this is the unique focal point only if $\rho^n \cdot a + (1-\rho^n)\cdot 0 \geq b \Leftrightarrow \rho \geq \sqrt[n]{\frac{b}{a}} = \rho_E$. Since ρ_E depends *positively* on *n* (observe that *b /a* <1) it follows that coordination on the E-equilibrium in the Stag Hunt becomes more complicated as *n* increases. This seems consistent with the evidence reported by Van Huyck et al. (1990) for the Weak-Link game and it is moreover rather intuitive: The efficient choice gets riskier as the number of players increase.

4. Conclusion and Extensions

This paper develops a theory based on the idea that the fairest equilibrium of a game –in case a unique one exists- is the obvious or prominent solution for many players, but also that players may disagree on what they deem fair, or alternatively that *some* players may not find obvious at all a fair equilibrium. We show the model to be consistent with abundant experimental evidence when we assume that some players consider efficiency *and* equity as necessary ingredients of fairness.

We argue that coordination between anonymous agents is relatively easy when it is common knowledge that ‘enough’ agents agree on the existence of a unique *fair* manner to play. On the contrary, coordination is delicate and might fail when there exist multiple fair equilibria or when there is just one fair equilibrium but the mass of players who agree on its obviousness is ‘small’.

We can mention a number of lessons from this. First, any team or group of people engaged in a joint task should make sure that most of them agree on a set of non-ambiguous normative values. Second, agents’ should be informed about other agents’ payoffs to facilitate the emergence of an obvious pattern of behavior. Third, teams should introduce additional coordination mechanisms like pre-play talk or the use of leaders when there exist multiple fair equilibria. Fourth, side payoffs conveniently altering either equilibrium or non-equilibrium payoffs may reduce the risk associated with play of the fair equilibrium and hence increase coordination, *even if they do not affect the set of equilibria*.

We believe that our model is more consistent with the available evidence on coordination that other selection criteria. Payoff dominance fails because it does not take into account that equity matters for making an equilibrium focal and because it is not consistent with the fact that out of equilibrium payoffs affect coordination. Risk dominance, although closely linked to our criterion, fails to take into account equity and efficiency.

One can think of a number of extensions which could be easily introduced. To start, an obvious one concerns extensive form games. Unfortunately, extending the model to this class of games is not direct. From our point of view, the main problem is introducing a parsimonious and at the same time empirically valid measure of the risk associated to playing the E-equilibrium. As another possible extension, one could relax the assumption of risk neutrality and assume instead prospect theory - Kahneman and Tversky (1979)- to investigate how negative payoffs may affect coordination.

Further, one might think of additional focal principles complementing the E-principle. For instance, our societies have conventions of the type 'first come, first served' that could help players to coordinate -for a particular example, see Sugden (1989). More precisely, we think of a focal principle which could be called the '*priority principle*' and that would complement hypothesis 3 when there exist multiple E-equilibria and an indisputably first (in a chronological sense) mover F: Any E-equilibrium that gives F the highest possible payoff is focal if the mass of E-types is large enough (in the sense of hypothesis 3).

We can illustrate this idea with one example. We noted before that our model was undetermined in the Battle of the Sexes (Game 11) because there were two E-equilibria. The priority principle does not help players to coordinate if they chose at the same time because there is no clear first mover in this case. In an interesting variation of the game, however, one player moves at time t whereas the other moves at time t+1 *without being informed of the other player's choice*. In this case, it is clear that the E-equilibrium that gives the highest payoff to the first mover is the unique focal point if ρ is larger than $b/(a + b)$. Hence, coordination seems more likely here than in the standard, 'truly simultaneous' version, a prediction supported by the evidence surveyed in Camerer (2003, pp. 365-7).

One important issue concerns communication. People sometimes believe that coordination *mainly* succeeds because agents are able to communicate their intentions and agree on a manner to do things. Although this is clearly an oversimplification,[16] abundant

[16] People often coordinate even if communication is unavailable (think of car drivers coordinating on which side of the road to drive in) or very costly –e.g., much time would be wasted if team workers

experimental evidence confirms that communication often fosters coordination on efficient equilibria –see, for instance, Cooper et al. (1992), and Duffy and Feltovich (2002). In a related work (…), we study this phenomenon.

To finish, our model predicts that an E-equilibrium is sure to be played if there is a critical mass of E-types, but it is undetermined otherwise. Our suggestion is that models of bounded rationality –or, more precisely, hybrid models combining sophisticated Nash players and agents with limited reasoning abilities- are likely candidates to explain behaviour in this case.

---

had to talk prior to starting *any* common effort. Further, communication raises a key question, that is, why people should believe others' announcements and express sincerely their intentions.

Bibliography

- Barry, B. (1995). *Justice as Impartiality*, University of California Press.
- Battalio, R., L. Samuelson, and J. Van Huyck (2001). "Optimization Incentives and Coordination Failure in Laboratory Stag Hunt Games", Econometrica, 69, 749-764.
- Brandts, J. and D. J. Cooper (forthcoming). "It is What You Say not What You Pay", forthcoming in the Journal of the European Economic Association.
- Camerer, C. (2003). *Behavioral Game Theory-Experiments in Strategic Interaction*, Princeton University Press.
- Clark, K., S. Kay, and M. Sefton (2001). "When are Nash Equilibria Self-Enforcing? An Experimental Analysis", International Journal of Game Theory 29, 495-515.
- Cooper, R., D. DeJong, R. Forsythe, and T. Ross (1989). "Communication in the Battle of the Sexes Game: Some Experimental Results", RAND Journal of Economics 20(4), 568-587.
- Cooper, R., D. DeJong, R. Forsythe and T. Ross (1990). "Selection Criteria in Coordination Games: Some Experimental Results", American Economic Review 80, 218-233.
- Cooper, R., D. DeJong, R. Forsythe and T. Ross (1992). "Communication in Coordination Games", Quarterly Journal of Economics 107, 739-771.
- Crawford, V. P., (1997). "Theory and Experiment in the Analysis of Strategic Interaction", in D. Kreps and K. Wallis (Eds.), *Advances in Economics and Econometrics: Theory and applications*, Seventh World Congress, Vol. I, Cambridge U.K.: Cambridge University Press.
- Devetag, G., and A. Ortmann (forthcoming). "When and Why? A Critical Survey on Coordination Failure in the Laboratory", forthcoming in Experimental Economics.
- Duffy, J., and N. Feltovich (2002). "Do Actions Speak Louder Than Words? Observation vs. Cheap Talk as Coordination Devices", Games and Economic Behavior, 39, 1-27.
- Fehr, E., and K. Schmidt (2006). "The Economics of Fairness, Reciprocity and Altruism – Experimental Evidence and New Theories", in S. C. Kolm, and J. M. Ythier (Eds.), *Handbook of the Economics of Giving, Altruism and Reciprocity*, Volume 1, Elsevier B. V.
- Gauthier, D. (1975). "Coordination", Dialogue, 14, 195-221.
- Harsanyi, J. C., and R. Selten (1988). *A General Theory of Equilibrium Selection in Games*, MIT Press.

- Haruvy, E., and D. O. Stahl (2007). "Equilibrium Selection and Bounded rationality in Symmetric Normal-Form Games", Journal of Economic Behavior and Organization, 62, 98-119.
- Kahneman, D., and A. Tversky (1979). "Prospect Theory: An Analysis of Decision under Risk", Econometrica, XLVII, 263-291.
- Keser, C. and B. Vogt (2000). "Why do experimental subjects choose an equilibrium which is neither risk nor payoff dominant?" Working Paper, University of Mannheim.
- Konow, J. (2003). "Which Is the Fairest One of All? A Positive Analysis of Justice Theories." Journal of Economic Literature, 41 (4), pp. 1186-1237.
- Mehta, J., C. Starmer, and R. Sugden (1994). "The nature of Salience: An Experimental Investigation of Pure Coordination Games", American Economic Review, 84, 658-673.
- Myerson, R. B., (1997). *Game Theory: Analysis of Conflict*. Cambridge: Harvard University Press (3rd edition).
- Roth, A. E., and M. W. Malouf (1979). "Game-Theoretic Models and the Role of Information in Bargaining", Psychological Review, 574-594.
- Schelling, T. (1960). *The Strategy of Conflict*. Cambridge: Harvard University Press.
- Schmidt, D., R. Shupp, J. M. Walker, and E. Ostrom (2003). "Playing Safe in Coordination Games: The Role of Risk Dominance, Payoff Dominance, Social History, and Reputation", Games and Economic Behavior, 42, 281-299.
- Stahl, D., and P., Wilson (1995). "On Players' Models of Other Players: Theory and Experimental Evidence", Games and Economic Behavior, 10, 213-254.
- Straub, P. G. (1995). "Risk Dominance and Coordination Failures in Static Games", The Quarterly Review of Economics and Finance, 35 (4), 339-363.
- Sugden, R. (1989). "Spontaneous Order", Journal of Economic Perspectives, 3(4), 85-97.
- Van Huyck, J. B., R. Battalio, and R. Beil (1990). "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failures", American Economic Review, 80, 234-248.
- Van Huyck, J. B., R. Battalio, and R. Beil (1991). "Strategic Uncertainty, Equilibrium Selection and Coordination Failure in Average Opinion Games", Quarterly Journal of Economics 106, 885-911.
- Van Huyck, J. B., A. Gillette, and R. Battalio (1992). "Credible Assignments in Coordination Games", Games and Economic Behavior, 4, 606-626.